

## CLAIMS

1. A method of identifying a consensus sequence for an intracellular antibody (ICS) comprising the steps of:
  - a) creating a database comprising sequences comprised by variable heavy chain domains and/or variable light chain domains of validated intracellular antibodies (VIDA database) and aligning the sequences of the variable heavy chain domains or variable light chain domains of validated intracellular antibodies, respectively;
  - b) determining the frequency with which a particular amino acid occurs in each of the positions of the aligned antibodies;
  - 10 c) selecting a frequency threshold value (LP or consensus threshold) in the range from 70% to 100%;
  - d) identifying the positions of the alignment at which the frequency of a particular amino acid is greater than or equal to the LP value;
  - e) identifying the most frequent amino acid, in the positions of the alignment defined in d),
  - 15 wherein a consensus sequence for an intracellular antibody is identified.
2. The method according to claim 1 wherein the sequences of the variable heavy chain domains or variable light chain domains of validated intracellular antibodies present in the VIDA database are aligned according to Kabat.
3. A method of identifying at least one optimum consensus sequence for an intracellular antibody (optimum ICS) comprising the steps of:
  - a) identifying a plurality of different ICSs for a plurality of different LP values;
  - b) for each of the plurality of different ICSs, constructing a frequency distribution of the number of identical amino acids between that particular ICS and each of the antibodies making up the VIDA database (VIDA distribution);
  - 20 c) for each of the plurality of ICSs, constructing a frequency distribution of the number of identical amino acids between that particular ICS and each of the antibodies that make up the Kabat database (Kabat distribution);
  - 25 d) defining a "distance" D between the VIDA distributions and the Kabat distribution corresponding to a value of LP;

- e) for each LP value, determining the value of the “distance” D between the VIDA distributions and the Kabat distribution corresponding to that value of LP;
- f) identifying the optimum ICS as the ICS corresponding to the value of LP for which the calculated value of the distance D defined in d) is maximum.

- 5 4. The method according to Claim 3 in which the consensus sequence is one of the consensus sequences for VH and/or VL comprising:
- a) for a VH consensus sequence, at least the following amino acids in the positions indicated according to the Chothia numbering:  
S-21, C-22, S-25, G-26, M-32, W-36, P-41, L-45, E-46, D-72, Q81, L-82c, E-85, D-86, A-88, Y-  
10 90, C-92, W-103, G-104, G-106, T-107, T-110, V-111, S-112;
  - b) for a VL consensus sequence, at least the following amino acids in the positions indicated according to the Chothia numbering:  
G-16, C-23, W-35, G-57, G-64, S-65, S-67, I-75, D-82, Y-86, C-88, T-102, K-103.

- 15 5. The method according to Claim 3 in which the consensus sequence is one of the consensus sequences for human VH and/or VL comprising:
- a) for human VH, essentially the following amino acids in the positions indicated according to the Chothia numbering:  
Q-1, V-2, Q-3, L-4, S-7, G-8, G-9, G10, V-12, P-14, G-15, S-17, L-18, R-19, L-20, S-21, C-22, A-24, S-25, G-26, F-27, T-28, F-29, Y-31a, M-32, W-36, R-38, Q-39, A-40, P-41, G-42, K-43,  
20 G-44, L-45, E-46, W-47, V-48, S-52, G-54, Y-58, Y-59, A-60, D-61, S-62, V-63, K-64, G-65, R-66, F-67, T-68, I-69, S-70, R-71, D-72, N-73, S-74, N-76, T-77, L-80, Q81, M-82, L-82c, R-83, A-84, E-85, D-86, T-87, A-88, Y-90, C-92, A-93, W-103, G-104, G-106, T-107, L-108, V-109, T-110, V-111, S-112, S-113;
  - b) for human VL, essentially the following amino acids in the positions indicated according to  
25 the Chothia numbering:  
T-5, P-8, G-16, I-21, C-23, W-35, Y-36, Q-37, P-40, G-41, P-44, I-48, S-56, G-57, S-63, G-64, S-65, S-67, G-68, L-73, T-74, I-75, D-82, A-84, Y-86, C-88, T-102, K-103.

6. A method of predicting whether an antibody is a functioning intracellular antibody, comprising the steps of:

- a) providing the sequence of an antibody molecule;
- b) aligning the sequence of the antibody with the sequences of the antibodies in a reference VIDA database;
- c) aligning the sequence of the antibody of step (a) with the optimum ICS sequence of said reference VIDA database;
- d) constructing the VIDA and KABAT distributions corresponding to the optimum value of LP (frequency threshold value);
- e) determining the corresponding distance D;
- f) determining the identity number N between the sequence of the antibody and the ICS reference sequence;
- g) calculating the difference between the mean value of the VIDA distribution and the product between D and the standard deviation of the VIDA distribution, obtaining the parameter  $S_{intra}$ ; and
- h) if the identity number N is greater than or equal to  $S_{intra}$ , said antibody molecule is predicted to be an intracellular antibody.

7. A method for conferring upon an immunoglobulin molecule the ability to function within an intracellular environment comprising the steps of:

- a) providing the sequence of an immunoglobulin molecule;
- b) identifying an optimum ICS reference sequence according to the method of claim 3; and
- c) modifying, by site-specific mutagenesis, at least one amino acid residue located in a position defined by said optimum ICS, in such a way that said amino acid residue is that identified by said optimum ICS.

8. The method of claim 7 wherein a plurality of amino acid residues are modified according to step (c).

- 9. The method of claim 7 wherein the amino acid residues that are located in the positions defined by said optimum ICS are modified according to step (c).

10. An intracellularly binding immunoglobulin molecule comprising a variable chain which is described by a consensus sequence selected from SEQ ID NO: 41 and SEQ ID NO: 42.
11. A method of selectively binding to a ligand in an intracellular environment, the method comprising contacting a molecule comprising a consensus sequence selected from the group  
5 consisting of SEQ ID NO: 41 and SEQ ID NO: 42, with a target ligand in an intracellular environment, wherein said target ligand is selectively bound by said molecule.
12. A method for identifying a consensus sequence for an intracellular antibody (ICS) comprising the steps of:
- (a) selecting and aligning a plurality of the sequences of antibody light or heavy chain variable  
10 regions which are shown using IACT to bind specifically to antigen or ligand within an intracellular environment; and
- (b) identifying the most frequent amino acid, in each position of the alignment, whereby, a consensus sequence is identified.
13. A method for selecting an antibody capable of binding specifically to one or more target  
15 antigen or ligand within an intracellular environment, comprising the steps of:
- (a) providing a plurality of antibodies that bind specifically to said one or more target antigens or ligands;
- (b) comparing at least a proportion of the variable heavy chain of one or more antibodies with consensus sequence SEQ ID NO: 3, and
- 20 (c) selecting an antibody whose variable light chain is at least 85% homologous with the consensus sequence of step (b), whereby an antibody that binds specifically to said one or more target antigens or ligands within an intracellular environment is selected.
14. An intracellularly binding immunoglobulin molecule comprising: a variable heavy chain which exhibits 85% homology to the consensus sequence SEQ ID No 3; and a variable light  
25 chain.

15. A method of selectively binding to a ligand in an intracellular environment, the method comprising contacting a molecule comprising:

a heavy chain domain (VH) that has at least 85% identity to the consensus sequence of SEQ ID NO: 3; and

5 at least one light chain domain (VL);

with a target ligand in an intracellular environment, wherein said target is selectively bound by said molecule.

16. A library, wherein the library is generated using any one or more of the variable heavy domain amino acid sequences (VH) selected from the group consisting of: a VH amino acid  
10 sequence showing at least 85% identity with the consensus sequence SEQ ID NO: 3 and a VH sequence according to SEQ ID No: 41.

17. A method of constructing an antibody library enriched with antibodies capable of functioning within an intracellular environment comprising the steps of:

a) selecting an antibody framework from those that are intracellularly functionally stable from  
15 the Kabat database;

b) on said framework, mutagenizing the amino acids present in the positions of the sequence defined by an optimum ICS, changing them into the amino acid residue that is located in that position in the optimum ICS sequence; and

c) on said framework, randomising the CDR regions of the antibody sequence, wherein a library  
20 enriched with antibodies that function within an intracellular environment is constructed.

18. A method of constructing an antibody library enriched with antibodies capable of functioning within an intracellular environment comprising the steps of:

a) selecting an initial antibody framework, on the basis of the maximum homology with an optimum ICS sequence;

b) mutagenizing all the remaining residues of the antibody framework, limited, for each position, to the amino acids that are located, in that position, in antibodies of the Kabat database;

c) on each of the frameworks, randomising the CDR regions of the antibody sequence, whereby an antibody library enriched with antibodies that function within an intracellular environment is constructed.

19. A method of producing immunoglobulin molecules, a substantial proportion of which selectively bind to a target ligand within an intracellular environment, the method comprising the steps of:

constructing an antibody library according to the method of claim 18;

10 contacting the members of said library with a plurality of target ligands in an intracellular environment, wherein a substantial proportion of said members selectively bind a member of said plurality of target ligands in said intracellular environment.